MathBio Journal Club:
Discussion of Stochastic Simulation

Will Heuett

University of Washington

Applied Mathematics Department

`heuett@amath.washington.edu`

21 October 2005

MathBio Journal Club:
Discussion of Stochastic Simulation

Will Heuett
University of Washington
Applied Mathematics Department
`heuett@amath.washington.edu`

21 October 2005

Abstract

We will discuss the uses of stochastic simulation modeling, follow Gillespie's derivation of his algorithm, and look at a few example biological systems.

Outline

★ Introduction: Where does stochastic simulation fit in?

Outline

★ Introduction: Where does stochastic simulation fit in?

★ Follow Gillespie's derivation of the algorithm.

Outline

★ Introduction: Where does stochastic simulation fit in?

★ Follow Gillespie's derivation of the algorithm.

★ Look at some examples.

Outline

★ Introduction: Where does stochastic simulation fit in?

★ Follow Gillespie's derivation of the algorithm.

★ Look at some examples.

★ Discuss what others have done to enhance the Gillespie algorithm to speed up computations or to include more information.

Outline

★ Introduction: Where does stochastic simulation fit in?

★ Follow Gillespie's derivation of the algorithm.

★ Look at some examples.

★ Discuss what others have done to enhance the Gillespie algorithm to speed up computations or to include more information.

★ Conclusions and Questions.

Outline

★ Introduction: Where does stochastic simulation fit in?

★ Follow Gillespie's derivation of the algorithm.

★ Look at some examples.

★ Discuss what others have done to enhance the Gillespie algorithm to speed up computations or to include more information.

★ Conclusions and Questions.

Introduction

"... biological systems are characterised by their regulatory and adaptive properties, from homeostatic mechanisms which maintain constant output levels to switching between alternative substrates or developmental pathways. Regulatory mechanisms including thresholds, allosteric interactions and feedback in gene transcription networks, metabolic pathways, signal transduction and intercellular interactions are defining biological characteristics - almost everything that happens in life boils down to enzyme-catalysed reactions."

- Crampin and Schnell, *Prog. Biophys. Mol. Biol.* (2004).

"... stochasticity is evident in all biological processes. The proliferation of both noise and noise reduction is a hallmark of organismal evolution."

- Fedoroff and Fontana, *Science* (2002).

Modeling Approaches
- Turner et al., Comp. Bio. Chem. (2004).

★ Directed graphs in which molecules are vertices and the reactions are the edges;

★ Bayesian networks in which the vertices correspond to random variables that describe, for example, a gene expression while the network defines a joint probability density function;

★ Boolean networks in which a biological object is either in an on or off state;

★ Ordinary differential equations (ODEs) in which chemical kinetics rate equations are used to represent protein concentrations;

★ Partial differential equations (PDEs) in which the spatial structure of cells are taken into account; and finally

★ Stochastic differential equations (SDEs) in which we have to resolve the issue of whether we work with concentrations or with individual molecules through continuous or discrete models.

Where does stochastic simulation fit in?

Three types of modeling regimes: discrete and stochastic, continuous and stochastic, and continuous and deterministic regimes.

★ Deterministic: the law of mass action.

★ Stochastic: the chemical master equation.

- Turner *et al.*, *Comp. Bio. Chem.* (2004).

Where does stochastic simulation fit in?

Three types of modeling regimes: discrete and stochastic, continuous and stochastic, and continuous and deterministic regimes.

★ Deterministic: the law of mass action.

★ Stochastic: the chemical master equation.

- Turner *et al.*, *Comp. Bio. Chem.* (2004).

Deterministic models are the infinite volume limit of the Markov chain models.

- Kurtz, *J. Chem. Phys.* (1972).

Outline

★ Introduction: Where does stochastic simulation fit in?

★ Follow Gillespie's derivation of the algorithm.

★ Look at some examples.

★ Discuss what others have done to enhance the Gillespie algorithm to speed up computations or to include more information.

★ Conclusions and Questions.

The Chemical Master Equation

The master equation is the time-evolution equation for the function $P(\mathbf{X}, t)$.

The Chemical Master Equation

The master equation is the time-evolution equation for the function $P(\mathbf{X}, t)$.

If there are M different reactions (events) then we have

$$P(\mathbf{X}, t+dt) = P(\mathbf{X}, t) \left[1 - \sum_{\mu=1}^{M} a_\mu dt \right] + \sum_{\mu=1}^{M} B_\mu dt,$$

which we rewrite as

$$\frac{\partial}{\partial t} P(\mathbf{X}, t) = \sum_{\mu=1}^{M} \left[B_\mu - a_\mu P(\mathbf{X}, t) \right]$$

where $a_\mu dt$ is the probability that reaction μ will occur in $(t, t+dt)$ given that the system is in state $\mathbf{X}$ at time t and $B_\mu dt$ is the probability that the system will go from being one reaction μ away from $\mathbf{X}$ at time t to $\mathbf{X}$ in $(t, t+dt)$.

Random Walk on an N-Dimensional Grid

So if we have a system with only 3 molecular species (R, L, and T), we model it as a Markov jump process where the system is doing a random walk on a 3-dimensional grid.
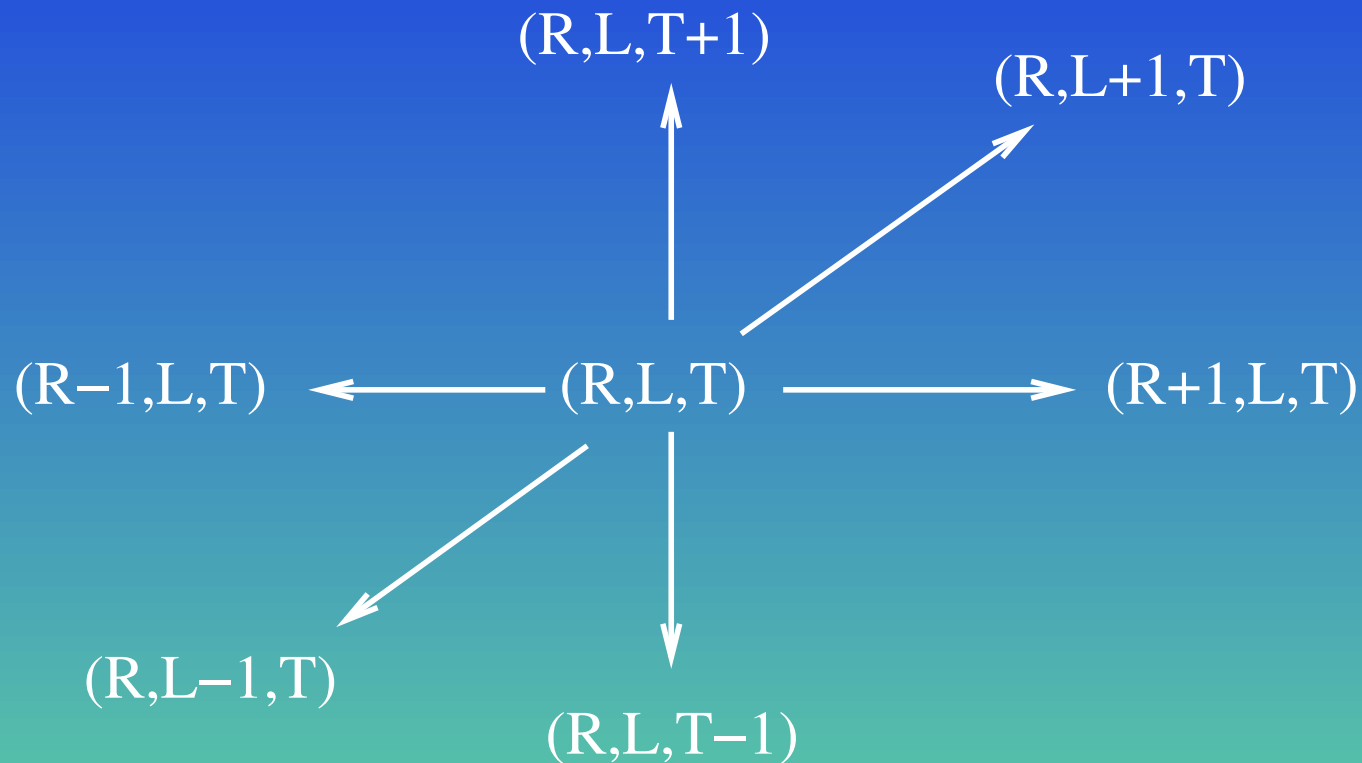
(R,L,T+1)

(R,L+1,T)

(R−1,L,T) ← (R,L,T) → (R+1,L,T)

(R,L−1,T)

(R,L,T−1)

The Reaction Probability Density Function

So we ask: "When will the next event occur and what type of event will it be?"

The Reaction Probability Density Function

So we ask: "When will the next event occur and what type of event will it be?"

Define

$$P(\tau, \mu)d\tau \equiv \text{probability at time } t \text{ that the next event will occur}$$

in the differential time interval $(t + \tau, t + \tau + d\tau)$

and will be a type μ event,

where $0 \leq \tau < \infty$ and μ simply indicates what type of event occurs.

The Reaction Probability Density Function

So we ask: "When will the next event occur and what type of event will it be?"

Define

$$P(\tau, \mu)d\tau \equiv \text{probability at time } t \text{ that the next event will occur}$$
$$\text{in the differential time interval } (t + \tau, t + \tau + d\tau)$$
$$\text{and will be a type } \mu \text{ event,}$$

where $0 \le \tau < \infty$ and μ simply indicates what type of event occurs.

This joint probability density function can be written as

$$P(\tau, \mu)d\tau = P_0(\tau)a_\mu d\tau$$

where $P_0(\tau)$ is the probability that no event occurs in the time interval $(t, t + \tau)$ and $a_\mu d\tau$ is the probability that event μ occurs in the interval $(t + \tau, t + \tau + d\tau)$.

The probability that no event occurs:

Let

$$a_0 = \sum_{i=1}^{M} a_i$$

The probability that no event occurs:

Let

$$a_0 = \sum_{i=1}^{M} a_i$$

then we have

$$P_0(t + dt) = P_0(t)(1 - a_0 dt)$$

The probability that no event occurs:

Let

$$a_0 = \sum_{i=1}^{M} a_i$$

then we have

$$P_0(t + dt) = P_0(t)(1 - a_0 dt)$$

or, rearranging a little, we have

$$\frac{P_0(t + dt) - P_0(t)}{dt} = -a_0 P_0(t)$$

The probability that no event occurs:

Let

$$a_0 = \sum_{i=1}^{M} a_i$$

then we have

$$P_0(t + dt) = P_0(t)(1 - a_0 dt)$$

or, rearranging a little, we have

$$\frac{P_0(t + dt) - P_0(t)}{dt} = -a_0 P_0(t)$$

from which it is easily deduced that

$$P_0(t) = e^{-a_0 t}.$$

The Reaction Probability Density Function

So we have

$$P(\tau, \mu)d\tau = P_0(\tau)a_\mu d\tau$$

The Reaction Probability Density Function

So we have

$$P(\tau, \mu)d\tau = P_0(\tau)a_\mu d\tau$$
$$= e^{-a_0\tau}a_\mu d\tau$$

The Reaction Probability Density Function

So we have

$$
\begin{aligned}
P(\tau, \mu)d\tau &= P_0(\tau)a_\mu d\tau \\
&= e^{-a_0\tau}a_\mu d\tau \\
&= a_0 e^{-a_0\tau}\left(\frac{a_\mu}{a_0}\right)d\tau
\end{aligned}
$$

The Reaction Probability Density Function

So we have

$$P(\tau, \mu)d\tau = P_0(\tau)a_\mu d\tau$$

$$= e^{-a_0\tau}a_\mu d\tau$$

$$= a_0 e^{-a_0\tau}\left(\frac{a_\mu}{a_0}\right)d\tau$$
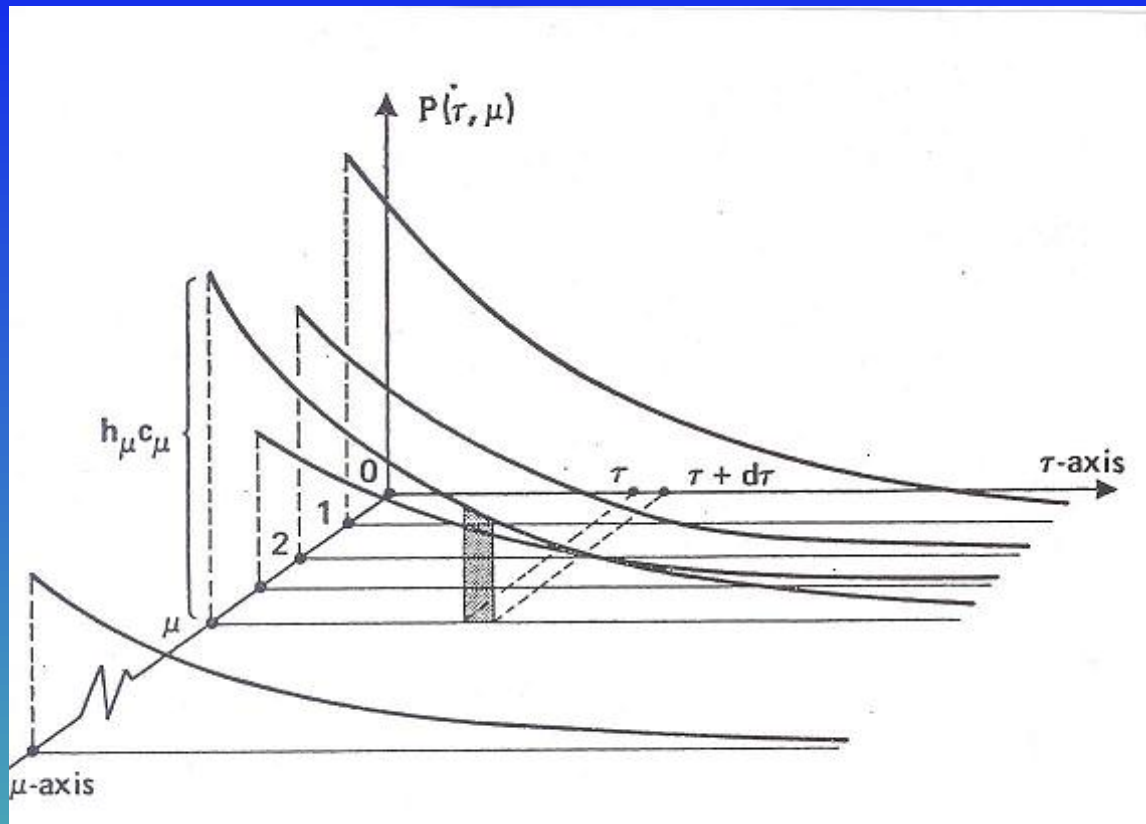
$$= P(\tau)P(\mu)d\tau$$

where

$$P(\mu) = \frac{a_\mu}{a_0}$$

and

$$P(\tau) = a_0 e^{-a_0\tau}.$$

Schematic of the Density Function

Gillespie, *J. Comp. Phys.* (1976).

The Probability Distribution Function

$$F(x) \equiv \int_{-\infty}^{x} P(x')dx'$$

The Probability Distribution Function

$$F(x) \equiv \int_{-\infty}^{x} P(x')dx'$$

To generate a random value x according to a given density function $P(x)$ we need to use the inversion method, by which we simply draw a random number r from the uniform distribution in the unit interval and take x such that

$$F(x) = r \quad \text{or} \quad x = F^{-1}(r)$$

The Probability Distribution Function

$$F(x) \equiv \int_{-\infty}^{x} P(x')dx'$$

To generate a random value x according to a given density function $P(x)$ we need to use the inversion method, by which we simply draw a random number r from the uniform distribution in the unit interval and take x such that

$$F(x) = r \quad \text{or} \quad x = F^{-1}(r)$$

since

$$F(x' + dx') - F(x') = F'(x')dx' = P(x')dx'.$$

Our Distribution Functions

$$P(\tau) = a_0 e^{-a_0 \tau} \quad \longrightarrow \quad F(\tau) = 1 - e^{-a_0 \tau}$$

$$P(\mu) = \frac{a_\mu}{a_0} \quad \longrightarrow \quad F(\mu) = \sum_{k=1}^{\mu} P(k)$$

Our Distribution Functions

$$P(\tau) = a_0 e^{-a_0 \tau} \quad \longrightarrow \quad F(\tau) = 1 - e^{-a_0 \tau}$$

$$P(\mu) = \frac{a_\mu}{a_0} \quad \longrightarrow \quad F(\mu) = \sum_{k=1}^{\mu} P(k)$$

So choose r_1 and r_2 from uniform distribution in the unit interval and

$$\tau = \frac{1}{a_0} \ln\left(\frac{1}{r_1}\right)$$

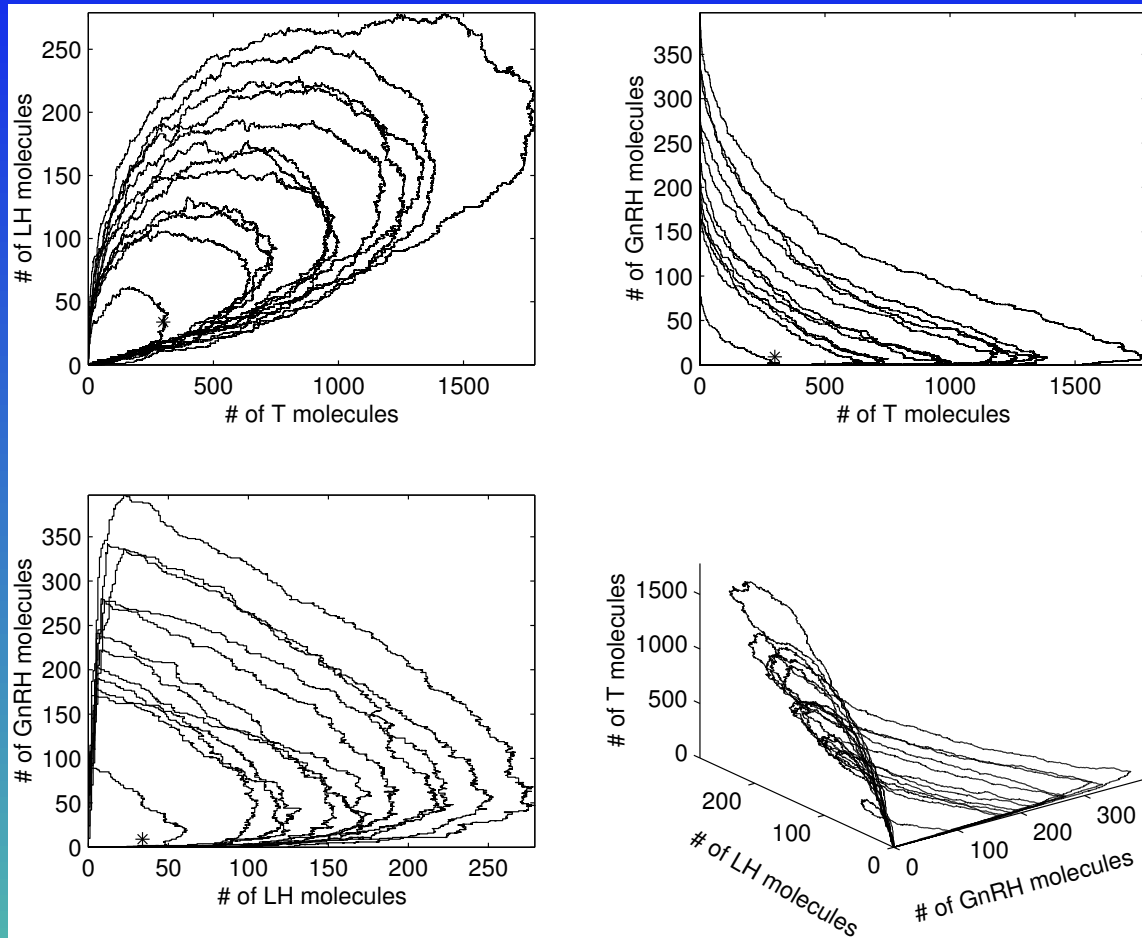$$\sum_{k=1}^{\mu-1} \frac{a_k}{a_0} < r_2 \leq \sum_{k=1}^{\mu} \frac{a_k}{a_0}.$$

Outline

★ Introduction: Where does stochastic simulation fit in?

★ Follow Gillespie's derivation of the algorithm.

★ Look at some examples.

★ Discuss what others have done to enhance the Gillespie algorithm to speed up computations or to include more information.

★ Conclusions and Questions.

A Stochastic Simulation
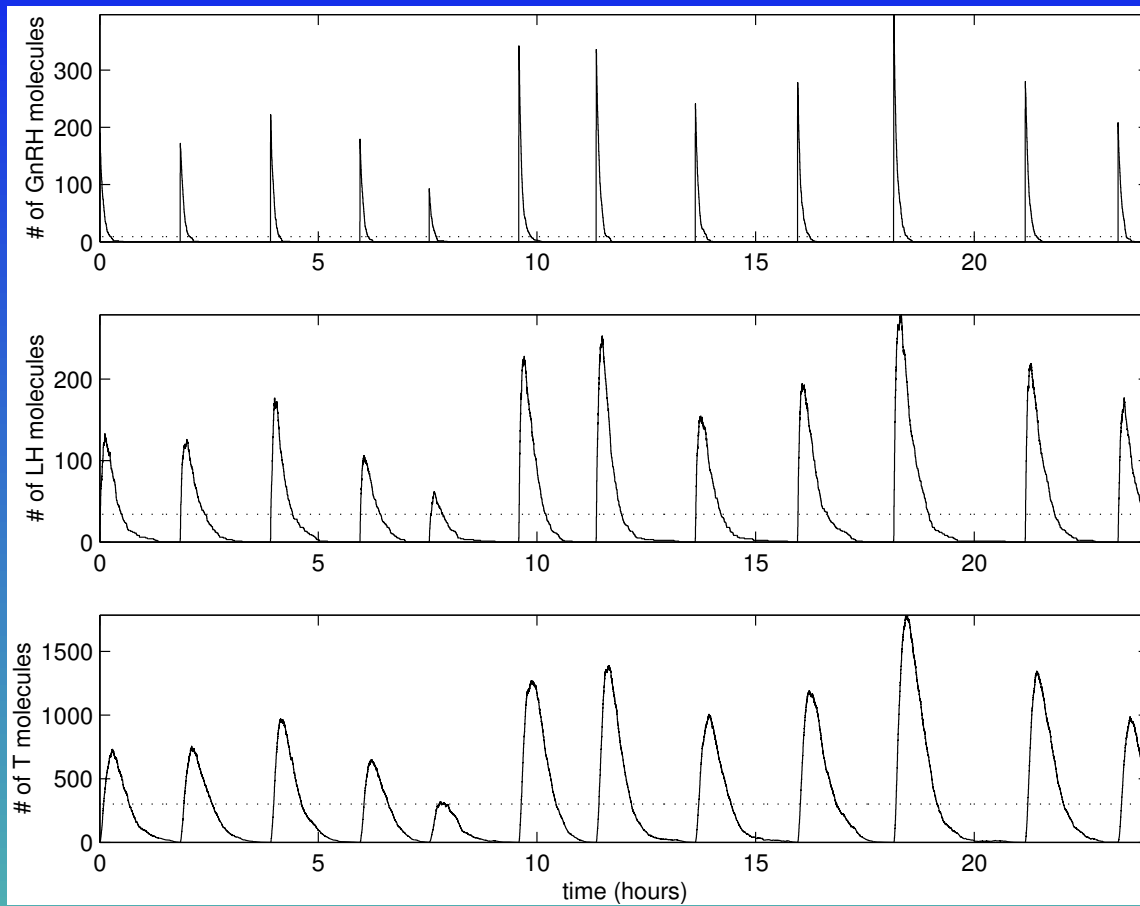
A Stochastic Simulation

Outline

⋆ Introduction: Where does stochastic simulation fit in?

⋆ Follow Gillespie's derivation of the algorithm.

⋆ Look at some examples.

⋆ Discuss what others have done to enhance the Gillespie algorithm to speed up computations or to include more information.

⋆ Conclusions and Questions.

Improvements...

★ Bortz, Kalos, and Lebowitz, *J. Comp. Phys.* (1975). - Speed up computations for Ising spin systems by accounting for priori probability of changing spins before, rather than after, choosing the spin or spins to change.

★ Turner, Schnell, and Burrage, *Comp. Bio. Chem.* (2004). - Include fluctuations caused by the structural organisation of the cytoplasm and the limited diffusion of molecules due to macromolecular crowding.

★ Burrage, Tian, and Burrage, *Prog. Biophys. Mol. Biol.* (2004). - Use multi-scale methods to incorporate the quasi-steady-state assumption with slow, intermediate, and fast reactions.

Conclusions

⋆ Stochastic effects are important in biological processes.

⋆ The Gillespie algorithm is a method for simulating stochastic processes.

⋆ This algorithm is easy to implement.

⋆ There may be ways to improve computational speeds or to incorporate additional stochastic effects for a particular system.

References

★ Bortz, A.B., Kalos, M.H., and Lebowitz, J.L. (1975) A new algorithm for Monte Carlo simulation of Ising spin systems. *J. Comp. Phys.* **17**, 10-18.

★ Burrage, K., Tian, T., and Burrage, P. (2004) A multi-scaled approach for simulating chemical reaction systems. *Prog. Biophys. Mol. Biol.* **85**, 217-234.

★ Gillespie, D.T. (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. of Comp. Phys.* **22**, 403-434.

★ Gillespie, D.T. (1977) Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340-2361.

★ Heuett, W.J. and Qian, H. (2005) A stochastic model of oscillatory blood testosterone levels. *Bull. Math. Bio.* submitted.

★ Kurtz, T.G. (1972) The relationship between stochastic and deterministic models for chemical reactions. *J. Chem. Phys.* **57**, 2976-2978.

★ Qian, H., Saffarian, S., and Elson, E.L. (2002) Concentration fluctuations in a mesoscopic oscillating chemical reaction system. *PNAS* **99**, 10376-10381.

★ Smolen, P., Baxter, D.A., and Byrne, J.H. (2000) Modeling transcriptional control in gene networks - methods, recent results, and future directions. *Bull. Math. Bio.* **62**, 247-292.

★ Turner, T.E., Schnell, S., and Burrage, K. (2004) Stochastic approaches for modelling in vivo reactions. *Comp. Bio. Chem.* **28**, 165-178.

★ Vilar, J.M.G., Kueh, H.Y., Barkai, N., and Leibler, S. (2002) Mechanisms of noise-resistance in genetic oscillators. *PNAS* **99**, 5988-5992.